Published in final edited form as:

Chem Biol Interact. 2013 February 25; 202(0): 22–31. doi:10.1016/j.cbi.2012.12.018.

ALDH16A1 is a novel non-catalytic enzyme that may be involved in the etiology of gout *via* protein–protein interactions with HPRT1

Vasilis Vasilou^{a,*}, Monica Sandoval^a, Donald S. Backos^a, Brian C. Jackson^a, Ying Chen^a, Philip Reigan^a, Miguel A. Lanasp^b, Richard J. Johnson^b, Vindhya Koppaka^a, and David C. Thompson^c

Vasilis Vasilou: vasilis.vasilou@ucdenver.edu

^aDepartment of Pharmaceutical Sciences, University of Colorado Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado, Aurora, CO 80045, USA

^bDivision of Nephrology, Department of Medicine, University of Colorado, Aurora, CO 80045, USA

^cDepartment of Clinical Pharmacy, University of Colorado Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado, Aurora, CO 80045, USA

Abstract

Gout, a common form of inflammatory arthritis, is strongly associated with elevated uric acid concentrations in the blood (hyperuricemia). A recent study in Icelanders identified a rare missense single nucleotide polymorphism (SNP) in the *ALDH16A1* gene, *ALDH16A1* *2, to be associated with gout and serum uric acid levels. *ALDH16A1* is a novel and rather unique member of the ALDH superfamily in relation to its gene and protein structures. *ALDH16* genes are present in fish, amphibians, protista, bacteria but absent from archaea, fungi and plants. In most mammalian species, two *ALDH16A1* spliced variants (*ALDH16A1*, long form and *ALDH16A1_v2*, short form) have been identified and both are expressed in HepG-2, HK-2 and HK-293 human cell lines. The *ALDH16* proteins contain two ALDH domains (as opposed to one in the other members of the superfamily), four transmembrane and one coiled-coil domains. The active site of *ALDH16* proteins from bacterial, frog and lower animals contain the catalytically important cysteine residue (Cys-302); this residue is absent from the mammalian and fish orthologs. Molecular modeling predicts that both the short and long forms of human *ALDH16A1* protein would lack catalytic activity but may interact with the hypoxanthine-guanine phosphoribosyltransferase (HPRT1) protein, a key enzyme involved in uric acid metabolism and gout. Interestingly, such protein-protein interactions with HPRT1 are predicted to be impaired for the long or short forms of *ALDH16A1**2. These results lead to the intriguing possibility that association between *ALDH16A1* and HPRT1 may be required for optimal HPRT activity with disruption of this interaction possibly contributing to the hyperuricemia seen in *ALDH16A1* *2 carriers.

© 2013 Elsevier Ireland Ltd. All rights reserved.

*Corresponding author. Address: Department of Pharmaceutical Sciences, University of Colorado Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado, 12850 E. Montview Blvd., Aurora, CO 80045, USA. Tel.: +1 303 724 3520; fax: +1 303 724 7266.

Conflict of interest statement: The authors declare no conflicts of interest associated with this manuscript.

Keywords

Aldehyde dehydrogenases; ALDH16A1; Gout; Hyperuricemia; HPRT1; Protein; protein interactions

1. Introduction

Gout is a painful inflammatory arthritis that more commonly afflicts men [29]. It is mediated by the extracellular deposition of urate crystals in joints. Not surprisingly, the major risk factor for gout is hyperuricemia, viz. uric acid levels >7 mg/dl in men and >6 mg/dl in women. Genome-wide searches have identified polymorphisms in genes (e.g., *ABCG2* and *SLC2A9*) as risk factors [28]; these account for approximately 6% of the variation of uric acid levels [40]. Most genetic associations to date have identified genes directly involved in urate transport in the kidney or intestine. However, a recent study from Iceland identified a rare missense single nucleotide polymorphism (SNP) in the *ALDH16A1* gene (*ALDH16A1**2) as a risk factor for hyperuricemia and gout [34]. *ALDH16A1* codes for a member of the aldehyde dehydrogenase (ALDH) superfamily [37]. Members of this superfamily oxidize a wide variety of aldehyde substrates and use NAD(P)⁺ as a coenzyme [21]. In addition to oxidizing aldehydes, some ALDHs exhibit other catalytic activities (e.g., esterase and reductase activities) and non-catalytic functions, such as structural and/or binding properties [21]. To date, no information on the catalytic properties of ALDH16A1 exists. A recent report on ALDH16A1 discussed its interaction with maspardin, a protein that, when truncated, is responsible for Mast syndrome [14]. In addition to maspardin, ALDH16A1 has been identified as a protein interacting with S-phase kinase-associated protein 1 (SKIP-1) [12], ubiquitin specific peptidase 1 (USP1) [32], proteasomal ATPase-associated factor 1 (PAAF1) [9] and protein kinase, AMP-activated, gamma 2 non-catalytic subunit (PRKAG2) [2]. Human ALDH16A1 has been found to interact with proteins from *Francisella tularensis*(*gltA*) and *Yersinia pestis*(*ymt*, *uxuA*, *icc*, and *fadR*), bacterial pathogens known to cause tularemia and the plague [7]. In addition, ALDH16A1 is predicted to interact with albumin (ALB), solute carrier family 2-facilitated glucose transporter (SLC2A4), heat shock protein 90 kDa alpha class B member 1 (cytosolic; HSP90AB1), (hypoxanthine phosphoribosyltransferase 1 (HPRT1), betaine–homocysteine S-methyltransferase (BHMT), and glioblastoma amplified sequence (GBAS) [17,22]. Of the various proteins listed, HPRT1 is a potentially important target for ALDH16A1 in the context of gout pathophysiology because HPRT1 has a key role in the purine salvage pathway. Mutations in the HPRT1 gene that diminish HPRT1 catalytic activity are the basis of a wide spectrum of clinical phenotypes, all of which are associated with significant overproduction of uric acid and related problems, such as hyperuricemia, urate nephrolithiasis, and gout. One of the mildest phenotypes, Kelley–Seegmiller syndrome, involves only problems related to overproduction of uric acid. The most severe phenotype, Lesch–Nyhan disease, involves a complete loss of HPRT1 catalytic activity and manifests as significant neurological dysfunction [16].

How a polymorphism in ALDH16A1, a protein assumed to be involved in aldehyde metabolism, could modulate uric acid levels is not immediately evident. Clearly, modification of uric acid metabolism and/or transport would be obvious mechanisms because these would lead to uric acid accumulation (Fig. 1). In this study, the human ALDH16A1 gene and protein, the alternative splice variants, and the evolution of the ALDH16 family of genes are described. In addition, molecular modeling is used to examine the catalytic activity of human ALDH16A1 protein (and its mutant ALDH16*2) and the feasibility of an interaction between these proteins and HPRT1.

2. Materials and methods

2.1. Evolutionary analysis of ALDH16 genes

ALDH16 genes were retrieved from the NCBI Gene database (<http://www.ncbi.nlm.nih.gov/gene>) using the terms 'ALDH16A1' or 'aldehyde dehydrogenase'. In addition, Blast analyses (<http://blast.ncbi.nlm.nih.gov/>) were used to retrieve genes with structure similar to the human *ALDH16A1*. Peptide sequences for each *ALDH16* gene were retrieved from the NCBI's Protein database (<http://www.ncbi.nlm.nih.gov/protein>) and aligned against a reference list of ALDH family members, including known human ALDHs and sequences from the NCBI's HomoloGene (<http://www.ncbi.nlm.nih.gov/homologene>), using the most accurate version of T-Coffee (<http://tcoffee.crg.cat>). Phylogenetic trees were constructed using a maximum likelihood method with 1000-replicate boot-strap in PHYLIP 3.69 (<http://evolution.genetics.washington.edu/phylip.html>). To be included for description, a gene record was required to meet two criteria: (i) the protein product of the gene must be 'full-length' (i.e., excludes fragments or partial records) and ii) the gene must have a known unique chromosomal location on the annotated genome.

2.2. Cell culture

Human cell lines examined included immortalized proximal tubule epithelial cell (HK-2), embryonic kidney cell (HEK-293), hepatoma cell (HepG-2), neuroblastoma cell (SH-SY5Y), retinal pigment epithelial cell (D407), adenocarcinomic alveolar basal epithelial cell (A549) and retinal pigment epithelial cell (ARPE-19) (ATCC, Manassas, VA). Cells were grown under standard conditions (37 °C, 5% CO₂ in air) and maintained in complete growth medium corresponding to each specific cell line (www.atcc.org). Cells were seeded onto 10 cm dishes and allowed to grow to ~80% confluence before harvesting for RNA extraction or whole cell lysate preparation.

2.3. Reverse transcription and PCR analysis

Total RNA from human cell lines (HK-2, HepG2 and HEK293) was extracted using Trizol (Invitrogen, Grand Island, NY) according to manufacturer's instructions. cDNA was reverse transcribed from total RNA (1 µg) using Maxima First strand cDNA synthesis kit with random hexamers (Fisher BioReagents, Fair Lawn, NJ). Primers specific for human *ALDH16A1* mRNA were designed and purchased from Integrated DNA Technologies (IDT, Coralville, IA). The forward and reverse primers were 5'-AGCCCATGGGAGTAATTGGCC-3' and 5'-AGACTCCTGGATGAGGAGCCTGAG-3', respectively. Semi-quantitative PCR was then conducted using GoTaq Green Master Mix (Promega, Madison, WI) according to manufacturer's instructions along with PCR control (no cDNA). Products were separated on a 2% agarose gel and visualized with ethidium bromide.

2.4. Western blot analysis

Whole cell lysates were prepared from the cell lines as previously described [5] and quantified for protein by Bradford method according to the manufacturer's protocol. Twenty µg of total protein was resolved on 10% SDS polyacrylamide gel, transferred onto polyvinylidene fluoride (PVDF) membrane (Millipore, Billerica, MA) and probed using an affinity purified rabbit polyclonal antibody (1:250) directed against the human *ALDH16A1* peptide sequence residues 307–324 (#5650; acetyl-CAIDPSMVSAEELEVQKGS-amide; BioSource International, Camarillo, CA; generous gift from Dr. Michael Hanna). After horseradish peroxidase-conjugated goat anti-rabbit IgG labeling, proteins were visualized using ECL-plus chemiluminescence reagent (Amersham, Pittsburgh, PA).

2.5. Computational-based molecular modeling

All molecular modeling studies were conducted using Accelrys Discovery Studio 3.1 (Accelrys Software, Inc., San Diego, CA). All crystal structure coordinates used in these studies were obtained from the protein data bank (<http://www.pdb.org>). The protein homology models of frog ALDH16A1 (*Xenopus laevis*) and both spliced variants of human ALDH16A1 were constructed by utilizing the structure of human mitochondrial ALDH [23] (PDB ID: 1O01; 45% sequence similarity with ALDH16A1) as a template using the MODELLER protocol [8]. Mutant protein structures of ALDH16A1*2 were made by changing the specific residue (Pro-527 or Pro-476) to Arg and both the wild-type and resulting mutant ALDH16A1 models were independently energy minimized prior to use in the *in silico* studies. Propionaldehyde and 4-hydroxynonenal were docked into human and frog ALDH16A1 using the flexible docking algorithm [18]. Potential protein–protein interactions of the N- and C-terminal domains of ALDH16A1, either with one another or with human HPRT1 [31] (PDB ID: 1BZY), were calculated using Zdock [6] with a rotation step of 6 degrees. Potential protein–protein complexes (54,000 total results) were scored based on shape complementarity as well as electrostatic and desolvation energies. The top 2000 results (based on ZDock score) were then subjected to additional energetic analysis using the ZRank algorithm [25], which represents a combination of van der Waals attractive and repulsive energies, short and long range repulsive and attractive energies, and desolvation. The top 20 potential protein–protein complexes (based on ZRank score) were then subjected to energy minimization and refinement with the RDock algorithm [19] and the structure and energy of the top result from each study reported. Loop refinement calculations of the flexible linker regions between the two domains were performed using the MODELLER protocol [8] to generate a series of 10 potential models that were scored with the Discrete Optimized Protein Energy (DOPE) function [30]; the highest scoring model was used in subsequent studies. A final energy minimization of the full complexes was then performed utilizing the conjugate gradient minimization protocol (Max steps: 10,000; RMS gradient: 0.1) with a CHARMM forcefield and the Generalized Born implicit solvent model with simple switching [11].

3. Results

3.1. Human ALDH16A1 gene structure and alternative splice variants

The human ALDH16A1 genes consist of 17 exons (Fig. 2A) encoding a 3119 bp mRNA and an 802 amino acid residue protein. Ten splice variants for ALDH16A1 are listed in the AceView database [35] and one of them has been included in the NCBI Gene database that we named as ALDH16A1_v2 (Fig. 2A). This variant encodes a 16 exon transcript (2966 bp mRNA and a 751 amino acid residue protein) that, to date, has been identified in human, chimpanzee, orang-utan, panda, pig, dog, guinea pig and hamster. The *ALDH16A1**2 allele carrying the missense SNP is located in exon 13 (c.1580C>G) of the 17-exon transcript and in exon 12 (c.1427C>G) of the 16-exon transcript (Fig. 2A). This leads to a missense proline to arginine alteration of amino acids 527 (p.Pro527Arg) and 476 (p.Pro476Arg) in NP_699160 and NP_001138868, respectively. Interestingly, the proline residue is conserved only in mammals whereas bacteria have arginine instead of proline at this position.

Two splice variants of the *ALDH16A1* gene have been identified in human cell lines, viz. *ALDH16A1* (17 exon) and *ALDH16A1_v2* (16 exon), with the length of the missing exon being 153bps. Semiquantitative RT-PCR of the variants in HK-2, HepG-2 and HEK-293 shows the ALDH16A1 variant to be more abundant than the *ALDH16A1_v2* variant (Fig. 2B).

Western blot analysis in a number of the human cell lines shows expression of ALDH16A1 protein at 75 kDa. A second band is detected at 55 kDa (Fig. 2C). To examine whether this

second band was caused by cross-reactivity with another ALDH protein, the ALDH16A1 antibody was tested against a number of recombinant ALDH proteins, including ALDH1A1, ALDH2, ALDH1B1, ALD-H3A1, and ALDH7A1. No cross-reactivity was found with these isozymes (data not shown).

3.2. ALDH16 genes

The ALDH16A1 gene appears to be conserved in mammalian species, including human, chimpanzee, dog, cow, mouse and rat. Our searches (as of July 2012) in all genome databases revealed ALDH16 genes to also be found in fish, amphibians, protista and bacteria but not in archaea, fungi and plants. The ALDH16 genes exhibit an interesting evolutionary pattern with at least 4 subfamilies, *viz.* 16A (mammals), 16B (amphibians and lower animals), 16C (bacteria) and 16D (fish) (Fig. 3). The amino acid identity within each of the subfamilies is more than 60% (data not shown).

3.3. ALDH16 proteins have a unique structure

While most members of the ALDH superfamily range between 400 and 570 aa in length and contain one ALDH domain [21], all ALDH16 proteins (from bacteria to mammals) are longer (750– 850 aa) and contain two ALDH protein domains: one within the first 494 aa and a second, shorter ALDH domain beginning at amino acid 525 (Fig. 4A). Four transmembrane domains and a coiled-coil domain are also predicted for the human ALDH16A1 protein (Fig. 4A).

3.4. Mammalian and fish ALDH16A1 proteins lack the “invariant” Cys-302 of the active site

Cys-302 is located in the active site of ALDH proteins and is considered essential for the catalytic activity of all ALDHs [3,26,38]. For example, in squid and octopus ALDH1C proteins, Arg substitutes for this invariant Cys-302; these proteins do not exhibit any ALDH enzymatic activity [41]. In the case of mammalian and fish ALDH16A1, the Cys-302 is missing (Fig. 4B) as a result of a gap in the sequence. This brings into question whether ALDH16A1 from these species would possess any catalytic activity. By contrast, Cys-302 is present in frog, sea anemone, squirt and lower animals, and in bacterial proteins (Fig. 4B), demonstrating that the absence of Cys-302 is not a property common to all ALDH16A1 proteins.

3.5. Lack of catalytic pocket in human ALDH16A1

In a space-filling homology model of frog ALDH16B1, the substrate and NAD⁺ binding pockets form tunnels that are clearly visible as white space in Fig. 5, and that converge with adequate room for a substrate and cofactor to interact. In addition, *in silico* substrate binding experiments indicate that the aldehydes, propional-dehyde and 4-hydroxynonenal, are each able to fit in the substrate-binding pocket and interact with the catalytic cysteine (Cys-291 in frog ALDH16B1 is homologous to Cys-302 in human ALDH3A1) and asparagine (Asn-160 in frog ALDH16B1 is homologous to Asn-169 in human ALDH3A1) (data not shown). These findings are consistent with structural studies of other ALDHs [20], and consistent with previous *in silico* substrate-binding studies performed by the authors. In contrast, the human ALDH16A1 homology model does not have clearly visible binding pockets for substrate or cofactor (NAD⁺). In addition, the tunnel between the substrate and cofactor binding sites that allows for their interaction is not present. When attempts were made to dock propionaldehyde or 4-hydroxynonenal into human ALDH16A1 *in silico*, no poses could be identified in which either substrate fit in the available space. Together, these data suggest that the frog ALDH16B1 has the necessary components to be catalytically-active as an ALDH, whereas human ALDH16A1 lacks both the specific amino acid residues and the substrate and cofactor binding pockets necessary for catalytic activity.

3.6. Molecular modeling of ALDH16A1–HPRT1 protein interactions

A homology model of human ALDH16A1 was generated utilizing human mitochondrial ALDH (PDB ID: 1O01) as the template for both the long and short forms of ALDH16A1. Both ALDH16A1 structures consist of a larger N-terminal domain and smaller C-terminal domain connected by a linker region that is approximately 30 residues in length (Fig. 6A and B). Loop refinement analyses of the linker regions of both forms of ALDH16A1 indicate this area is highly likely to be disordered and, therefore, flexible. This introduces the possibility that ALDH16A1 could exist as either an “open” form with the two domains separated, or a “closed” form in which the two domains are associated together. To examine this possibility, computational modeling of the predicted protein-protein interactions between the N- and C-terminal domains of the protein homology models was performed. These studies indicated that interaction between the two domains was feasible, with favorable predicted interaction energies occurring between the two domains for both the long and short forms. Representative structures of the closed conformations and their associated interaction energies are depicted in Fig. 6C and D.

To assess the effect of the respective Pro to Arg mutations (i.e., P527R for the long and P476R for the short) on formation of the closed form of ALDH16A1, the predicted interactions between the N- and C-terminal domains of the mutant structures were examined. These studies indicated no difference in the predicted interaction energies between the two domains in response to the mutation (data not shown). The stability of the protein complexes of the Pro to Arg mutants with the wild-type was also compared by energy minimization of the resulting structures. The mutation in the long form (P527R) lowered the energy of the closed form complex relative to the wild-type (WT) form (P527R: -35739.2 kcal/mol vs. WT: -35665.6 kcal/mol). Similar results were obtained for P476R mutation in the short form (P476R: -32807.8 kcal/mol vs. WT: -32729.0 kcal/mol). This suggests that the closed conformations of both the ALDH16A1 mutants may be more stable than their wild-type counterparts.

Given that ALDH16A1 has been shown to interact with a variety of other cellular proteins, the potential for interaction between the closed or open conformations of both the long and short forms of ALDH16A1 with HPRT1 were examined *in silico*. These studies also included determination of the predicted effects of the P527R/P476R mutations, which are located at the start of the C-terminal domain, on either the protein-protein interactions or the relative stability of any resulting protein complexes. No likely protein-protein interactions were identified between HPRT1 and the closed conformations of either form of wild-type ALDH16A1, suggesting a low probability for interaction with HPRT1. Predicted interactions between the N- and C-terminal domains of the open conformations for both wild-type forms were much more favorable (Fig. 7A and B). The P527R/P476R mutations did result in a slight difference in the predicted energy of interaction between the C-terminal domain and HPRT1 relative to their respective wild-type form (Table 1). Energy minimization was employed to predict the relative stability of the wild-type and mutant ALDH16A1-HPRT1 complexes. In both ALDH16A1 spliced variants, the mutation resulted in a lower predicted stability (i.e., higher relative energy) than its respective wild-type counterpart (Table 2). Although the predicted interaction of the long form (Fig. 7A) appears to involve opposing subunits of HPRT1 (as compared with the short form (Fig. 7B), the homotetrameric nature of HPRT1 means that either orientation would be just as likely for both spliced variants.

4. Discussion

This study describes the structure, evolution and functions of human ALDH16A1 gene and protein. The ALDH16 genes and proteins are interesting members of the ALDH gene

superfamily for many reasons. The *ALDH16A1* gene is highly conserved in mammalian species and two *ALDH16A1* spliced variants have been identified in humans (*ALDH16A1*, long form and *ALDH16A1_v2*, short form) and in chimpanzees, orangutans, pandas, pigs, dogs, guinea pigs and hamsters. Review of the AceView database indicated that the number of human mRNA/cDNA clones deposited in public repositories (i.e., GenBank and dbEST) includes 37 clones for the long form (*ALDH16A1*) and 139 for the short form (*ALDH16A1_v2*). Both of these variants were transcribed in three of the human cell lines (HepG-2, HK-2 and HK293) examined in this study. All three cell lines expressed *ALDH16A1*, a 75 kDa protein. A second band at 55 kDa was also detected by the anti-human *ALDH16A1* antibody. This may have been the result of a cross reactivity with another ALDH protein. However, no cross-reactivity of the *ALDH16A1* anti-body was observed with any of the ALDH recombinant proteins tested, i.e., *ALDH1A1*, *ALDH2*, *ALDH1B1*, *ALDH3A1* and *ALDH7A1* (data not shown). Another interesting observation was that *ALDH16* genes are found in fish, amphibians, protista, bacteria but not in archaea, fungi and plants. The reason for the absence of this gene from these species is currently unknown.

All *ALDH16* proteins from bacteria to mammals contain two ALDH protein domains. This distinguishes *ALDH16* from all other ALDHs, which contain only one protein domain. The four transmembrane domains found in human *ALDH16A1* are consistent with it being membrane-associated. Supporting this contention are our preliminary data on the expression of this protein in human and mouse tissues that confirms the localization of *ALDH16* to the cytosol and membranes (data not shown). ALDHs were originally named as a result of their enzymatic activity. The amino acid residue Cys-302 in the ALDH active site is considered critical to ALDH catalytic activity [10]. In the present study, fish and mammalian *ALDH16* proteins were identified as lacking Cys-302. Based upon the results of previous studies, the *ALDH16* proteins from these species would be expected to lack catalytic activity. *In silico* comparison of the active site structure of human *ALDH16A1* with that of frog *ALDH16* (which contains the equivalent of Cys-302) revealed substrate and cofactor binding tunnels in the frog protein that would allow substrate and cofactor interaction. By contrast, no such tunnels are apparent in the human protein, suggesting that such a structural change in human *ALDH16A1* would hinder appropriate substrate orientation in the active site and thereby prevent substrate catalysis. This was verified *in silico* when docking poses for aldehyde substrates could be found in frog *ALDH16* but not in human *ALDH16A1*. Closer examination of the *ALDH16* family of genes uncovered an interesting evolutionary pattern that correlates with the catalytic properties (i.e., lack or presence of the Cys-302) of their gene products. Specifically, the 16A (mammalian proteins) and 16D (fish proteins) subfamilies do not have Cys-302, whereas the 16B (amphibians and lower animals) and 16C (bacteria) subfamilies contain Cys302. The loss of Cys-302 in mammals and fish and the associated disruption of the catalytic pocket may have resulted from an evolutionary selective pressure that caused inactivation of these enzymes, potentially to allow them to perform some other non-catalytic function.

The presence of catalytically-inactive homologues in families of enzymes has become increasingly appreciated and confirmed by the large scale sequencing of many genomes. The existence of so-called “inactive enzyme-homologues” [27], “nonenzymes” [36] or “dead enzymes” [1] was first confirmed in 1967 with the discovery of α -lactalbumin, a non-enzymatic form of lysozyme [4]. Since then, it has become clear that inactive enzymes are present in many enzyme families, such as proteases, phosphatases, kinases, phospholipases, just to name a few. Lens crystallins, described in 1987, are considered to be the result of gene sharing, *viz.* the recruitment of enzymes as structural elements of the eye [24,39]. Among these proteins are the η - and ω -lens crystallins which have been identified as ALDHs, specifically *ALDH1A8* [13], *ALDH1A9* [15] and *ALDH1C* [41]. Based on the

results obtained in the present study, the ALDH16 proteins in mammals and fish seem likely to be included in the “nonenzyme” classification.

ALDH16A1 consists of two ALDH domains separated by a flexible linker region. This is a unique structural feature (i.e., not shared with the other members of the ALDH super family) that allows for the possibility of interaction between the two domains. The interaction interface between the N- and C-terminal domains of both spliced variants of ALDH16A1 are also the regions predicted to interact with HPRT1. It is therefore not surprising that the closed forms of ALDH16A appear to abolish the capacity of ALDH16A1 to associate with HPRT1. The observation that the P527R/P476R (i.e., Pro to Arg) mutation resulted in a modest change in the predicted interaction energy between the N- and C-terminal domains of either spliced variant would suggest that the mutation does not necessarily alter the propensity for interaction between the two domains *per se*. However, the predicted stability of the closed form complexes was substantially higher in the mutant proteins than in the wild-type proteins, suggesting that the mutation favors the continued association between the N- and C-terminals after formation of the closed structure. The mutation may thus lead to greater levels of the closed conformation relative to the open conformation. The Pro to Arg mutations only had a minor impact on the predicted interaction of the C-terminal domains of ALDH16A1 with HPRT1. Nevertheless, the complex formed with HPRT1 and either mutant protein was predicted to be less stable than that formed with the corresponding wild-type protein. Taken together, these data suggest that the overall effect of the P527R/P476R mutation would be to impair or prevent the association between ALDH16A1 and HPRT1 by favoring the formation of the closed form of ALDH16A1 (which was not predicted to interact with HPRT1) and/or by impeding the formation and maintained association of the ALDH16A1–HPRT1 complex.

The structure of HPRT1 is homotetrameric with the two diagonally opposed subunits in a position parallel to each other and antiparallel to the two remaining subunits. Given that the opposing surfaces of HPRT1 on all three axes (*x*, *y*, and *z*) are mirror images of each other, the predicted interactions of HPRT1 with either the long or short forms of ALDH16A1 in the opposite orientation (i.e., mirrored across the midline of HPRT1) would be equally valid. Thus, it is possible that HPRT1 could interact with two ALDH16A1s simultaneously. Furthermore, the binding orientations of the long versus short spliced variant forms of ALDH16A1 would not preclude the possibility of HPRT1 interacting simultaneously with one ALDH16A1 long and one ALDH16A1 short. It is not possible to predict at the present time whether these *in silico*-predicted interactions with HPRT1 occur *in vitro* or *in vivo*. Certainly, if these associations do indeed occur within the cell, it is intriguing to speculate whether the various permutations in ALDH16A1 binding partners (e.g., one long, two long, one short, two short or one of each) represent a means by which HPRT1 activity could be subject to multi-level regulation by ALDH16A1.

It is currently unclear what functional effects the association between HPRT1 and ALDH16A1 may have on HPRT1 activity. The predicted binding of these two proteins does not physically occlude the active sites of HPRT1, making it unlikely that the association would induce inhibition by interfering with substrate binding. However, we did not observe substantial changes in the volume or conformation of these sites in response to association with ALDH16A1 that would indicate ALDH16A1 enhances HPRT1 activity *via* allosteric modification. In any event, the present results indicate the feasibility of a possible association between ALDH16A1 and HPRT1 that is predicted to be impaired by the P527R/P476R mutation. Given that this mutation has been associated with a high incidence of gout, a disease caused by a dysfunction in HPRT1, it would be reasonable to assume the interaction with wild-type ALDH16A1, if it does occur within the cellular environment, would have a functionally positive effect on HPRT1 activity. We are currently investigating

whether association with ALDH16A1 alters the predicted binding of substrates to HPRT1 *in silico*.

Acknowledgments

We would like to thank Dr. Michael Hanna (Texas A & M University) for providing the ALDH16A1 antibody and also the Computational Chemistry and Biology Core Facility at the University of Colorado Anschutz Medical Campus for their contributions to these studies. This work was supported, in part, by the following NIH grants; EY17963 (V.V.) EY11490 (V.V.) and F31 AA020728 (B.C.J).

References

1. Adrain C, Freeman M. New lives for old: evolution of pseudoenzyme function illustrated by iRhoms. *Nat Rev Mol Cell Biol.* 2012; 13:489–498. [PubMed: 22781900]
2. Behrends C, Sowa ME, Gygi SP, Harper JW. Network organization of the human autophagy system. *Nature.* 2010; 466:68–76. [PubMed: 20562859]
3. Blatter EE, Abriola DP, Pietruszko R. Aldehyde dehydrogenase. Covalent intermediate in aldehyde dehydrogenation and ester hydrolysis. *Biochem J.* 1992; 282(Pt 2):353–360. [PubMed: 1546951]
4. Brew K, Vanaman TC, Hill RL. Comparison of the amino acid sequence of bovine alpha-lactalbumin and hens egg white lysozyme. *J Biol Chem.* 1967; 242:3747–3749. [PubMed: 6038502]
5. Brocker C, Cantore M, Failli P, Vasiliou V. Aldehyde dehydrogenase 7A1 (ALDH7A1) attenuates reactive aldehyde and oxidative stress induced cytotoxicity. *Chem Biol Interact.* 2011; 191:269–277. [PubMed: 21338592]
6. Chen R, Weng Z. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins.* 2002; 47:281–294. [PubMed: 11948782]
7. Dyer MD, Neff C, Dufford M, Rivera CG, Shattuck D, Bassaganya-Riera J, Murali TM, Sobral BW. The human–bacterial pathogen protein interaction networks of *Bacillus anthracis*, *Francisella tularensis*, and *Yersinia pestis*. *PLoS One.* 2010; 5:e12089. [PubMed: 20711500]
8. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. Comparative protein structure modeling using Modeller. *Curr Protoc Bioinf.* 2006; Chapter 5, Unit 5.6
9. Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, McBroom-Cerajewski L, Robinson MD, O'Connor L, Li M, Taylor R, Dharsee M, Ho Y, Heilbut A, Moore L, Zhang S, Ornatsky O, Bukhman YV, Ethier M, Sheng Y, Vasilescu J, Abu-Farha M, Lambert JP, Duewel HS, Stewart, Kuehl B, Hogue K, Colwill K, Gladwish K, Muskat B, Kinach R, Adams SL, Moran MF, Morin GB, Topaloglou T, Figeys D. Large-scale mapping of human protein–protein interactions by mass spectrometry. *Mol Syst Biol.* 2007; 3:89. [PubMed: 17353931]
10. Farres J, Wang TT, Cunningham SJ, Weiner H. Investigation of the active site cysteine residue of rat liver mitochondrial aldehyde dehydrogenase by site-directed mutagenesis. *Biochemistry.* 1995; 34:2592–2598. [PubMed: 7873540]
11. Feig M, Onufriev A, Lee MS, Im W, Case DA, Brooks CL 3rd. Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *J Comput Chem.* 2004; 25:265–284. [PubMed: 14648625]
12. Foster LJ, Rudich A, Talior I, Patel N, Huang X, Furtado LM, Bilan PJ, Mann M, Klip A. Insulin-dependent interactions of proteins with GLUT4 revealed through stable isotope labeling by amino acids in cell culture (SILAC). *J Proteome Res.* 2006; 5:64–75. [PubMed: 16396496]
13. Graham C, Hodin J, Wistow G. A retinaldehyde dehydrogenase as a structural protein in a mammalian eye lens. Gene recruitment of eta-crystallin. *J Biol Chem.* 1996; 271:15623–15628. [PubMed: 8663049]
14. Hanna MC, Blackstone C. Interaction of the SPG21 protein ACP33/masparidin with the aldehyde dehydrogenase ALDH16A1. *Neurogenetics.* 2009; 10:217–228. [PubMed: 19184135]
15. Horwitz J, Ding L, Vasiliou V, Cantore M, Piatigorsky J. Scallop lens Omega-crystallin (ALDH1A9): a novel tetrameric aldehyde dehydrogenase. *Biochem Biophys Res Commun.* 2006; 348:1302–1309. [PubMed: 16919242]

16. Kelley WN, Rosenbloom FM, Henderson JF, Seegmiller JE. A specific enzyme defect in gout associated with overproduction of uric acid. *Proc Natl Acad Sci USA*. 1967; 57:1735–1739. [PubMed: 4291947]
17. Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, Duesbury M, Dumousseau M, Feuermann M, Hinz U, Jandrasits C, Jimenez RC, Khadake J, Mahadevan U, Masson P, Pedruzzi I, Pfeifferberger E, Porras P, Raghunath A, Roechert B, Orchard S, Hermjakob H. The IntAct molecular interaction database in 2012. *Nucleic Acids Res*. 2012; 40:D841–D846. [PubMed: 22121220]
18. Koska J, Spassov VZ, Maynard AJ, Yan L, Austin N, Flook PK, Venkatachalam CM. Fully automated molecular mechanics based induced fit protein-ligand docking method. *J Chem Inf Model*. 2008; 48:1965–1973. [PubMed: 18816046]
19. Li L, Chen R, Weng Z. RDOCK: refinement of rigid-body protein docking predictions. *Proteins*. 2003; 53:693–707. [PubMed: 14579360]
20. Liu ZJ, Sun YJ, Rose J, Chung YJ, Hsiao CD, Chang WR, Kuo I, Perozich J, Lindahl R, Hempel J, Wang BC. The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold. *Nat Struct Biol*. 1997; 4:317–326. [PubMed: 9095201]
21. Marchitti SA, Bocker C, Stagos D, Vasilidou V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. *Expert Opin Drug Metab Toxicol*. 2008; 4:697–720. [PubMed: 18611112]
22. Niu Y, Otasek D, Jurisica I. Evaluation of linguistic features useful in extraction of interactions from PubMed; application to annotating known, high-throughput and predicted interactions in I2D. *Bioinformatics*. 2010; 26:111–119. [PubMed: 19850753]
23. Perez-Miller SJ, Hurley TD. Coenzyme isomerization is integral to catalysis in aldehyde dehydrogenase. *Biochemistry*. 2003; 42:7100–7109. [PubMed: 12795606]
24. Piatigorsky J, Wistow GJ. Enzyme/crystallins: gene sharing as an evolutionary strategy. *Cell*. 1989; 57:197–199. [PubMed: 2649248]
25. Pierce B, Weng Z. ZRANK: reranking protein docking predictions with an optimized energy function. *Proteins*. 2007; 67:1078–1086. [PubMed: 17373710]
26. Pietruszko R, Blatter E, Abriola DP, Prestwich G. Localization of cysteine 302 at the active site of aldehyde dehydrogenase. *Adv Exp Med Biol*. 1991; 284:19–30. [PubMed: 2053476]
27. Pils B, Schultz J. Inactive enzyme-homologues find new function in regulatory processes. *J Mol Biol*. 2004; 340:399–404. [PubMed: 15210342]
28. Riches PL, Wright AF, Ralston SH. Recent insights into the pathogenesis of hyperuricaemia and gout. *Hum Mol Genet*. 2009; 18:R177–184. [PubMed: 19808794]
29. Rott KT, Agudelo CA. Gout. *JAMA*. 2003; 289:2857–2860. [PubMed: 12783917]
30. Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. *Protein Sci*. 2006; 15:2507–2524. [PubMed: 17075131]
31. Shi W, Li CM, Tyler PC, Furneaux RH, Grubmeyer C, Schramm VL, Almo SC. The 2.0 Å structure of human hypoxanthine–guanine phosphoribosyltransferase in complex with a transition-state analog inhibitor. *Nat Struct Biol*. 1999; 6:588–593. [PubMed: 10360366]
32. Sowa ME, Bennett EJ, Gygi SP, Harper JW. Defining the human deubiquitinating enzyme interaction landscape. *Cell*. 2009; 138:389–403. [PubMed: 19615732]
33. Steinmetz CG, Xie P, Weiner H, Hurley TD. Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion. *Structure*. 1997; 5:701–711. [PubMed: 9195888]
34. Sulem P, Gudbjartsson DF, Walters GB, Helgadóttir HT, Helgason A, Gudjonsson SA, Zanon C, Besenbacher S, Björnsdóttir G, Magnusson OT, Magnusson G, Hjartarson E, Saemundsdóttir J, Gylfason A, Jonasdóttir A, Holm H, Karason A, Rafnar T, Stefansson H, Andreassen OA, Pedersen JH, Pack AI, de Visser MC, Kiemeny LA, Geirsson AJ, Eyjólfsson GI, Olafsson I, Kong A, Masson G, Jonsson H, Thorsteinsdóttir U, Jónsdóttir I, Stefansson K. Identification of low-frequency variants associated with gout and serum uric acid levels. *Nat Genet*. 2011; 43:1127–1130. [PubMed: 21983786]
35. Thierry-Mieg D, Thierry-Mieg J. AceView: a comprehensive cDNA-supported gene and transcripts annotation. *Genome Biol*. 2006; 7(Suppl. 1):S12.11–S12.14. [PubMed: 16925834]

36. Todd AE, Orengo CA, Thornton JM. Sequence and structural differences between enzyme and nonenzyme homologs. *Structure*. 2002; 10:1435–1451. [PubMed: 12377129]
37. Vasiliou V, Nebert DW. Analysis and update of the human aldehyde dehydrogenase (ALDH) gene family. *Hum Genomics*. 2005; 2:138–143. [PubMed: 16004729]
38. Weiner H, Farres J, Wang TT, Cunningham SJ, Zheng CF, Ghenbot G. Probing the active site of aldehyde dehydrogenase by site directed mutagenesis. *Adv Exp Med Biol*. 1991; 284:13–17. [PubMed: 2053473]
39. Wistow G, Piatigorsky J. Recruitment of enzymes as lens structural proteins. *Science*. 1987; 236:1554–1556. [PubMed: 3589669]
40. Yang Q, Kottgen A, Dehghan A, Smith AV, Glazer NL, Chen MH, Chasman DI, Aspelund T, Eiriksdottir G, Harris TB, Launer L, Nalls M, Hernandez D, Arking DE, Boerwinkle E, Grove ML, Li M, Linda Kao WH, Chonchol M, Haritunians T, Li G, Lumley T, Psaty BM, Shlipak M, Hwang SJ, Larson MG, O'Donnell CJ, Upadhyay A, van Duijn CM, Hofman A, Rivadeneira F, Stricker B, Uitterlinden AG, Pare G, Parker AN, Ridker PM, Siscovick DS, Gudnason V, Witteman JC, Fox CS, Coresh J. Multiple genetic loci influence serum urate levels and their relationship with gout and cardiovascular disease risk factors. *Circ Cardiovasc Genet*. 2010; 3:523–530. [PubMed: 20884846]
41. Zinovieva RD, Tomarev SI, Piatigorsky J. Aldehyde dehydrogenase-derived omega-crystallins of squid and octopus. Specialization for lens expression. *J Biol Chem*. 1993; 268:11449–11455. [PubMed: 7684383]

Abbreviations

ALDH	aldehyde dehydrogenase
HPRT	hypoxanthine–guanine phosphoribosyltransferase
SNP	single nucleotide polymorphism
SKIP-1	S-phase kinase-associated protein 1
PAAF	proteasomal ATPase-associated factor
PRKAG2	protein kinase, AMP-activated, gamma 2 non-catalytic subunit
ALB	albumin
SLC2A4	solute carrier family 2-facilitated transporter
HSP90	heat shock protein 90 kDa
BHMT	Betaine–homocysteine S-methyltransferase
GBAS	glioblastoma amplified sequence
PVDF	polyvinylidene fluoride
DOPE	discrete optimized protein energy
AMP	adenosine monophosphate
IMP	inosine monophosphate
AMPD	adenosine monophosphate deaminase
PNP	purine nucleoside phosphorylase
URAT	uric acid transporter
OAT	organic anion transporter
glut	glucose transporter
ABC	ATP-binding cassette

MRP	multidrug resistance protein
NPT	sodium-phosphate transporter
TM	trans-membrane domains
C-C	coiled-coil domain

Fig. 1.

Production and transport of uric acid in epithelial cells. Uric acid is the final product of adenosine monophosphate (AMP) catabolism. Initially, AMP can be phosphorylated back to ADP and ATP by adenylate kinase and also can stimulate ATP production by stimulation of AMP kinase, glycolysis and fat oxidation. Alternatively, AMP can enter in the purine degradation pathway which is initialized by its deamination to inosine monophosphate (IMP) by adenosine monophosphate deaminase (AMPD). Elimination of the ribose from IMP by 5⁰nucleotidase forms inosine which is then metabolized to hypoxanthine by purine nucleoside phosphorylase (PNP). Inosine can also be produced from the deamination of adenosine by adenosine deaminase. Alternatively, adenosine can also enter the route through its phosphorylation to AMP by adenosine kinase. Hypoxanthine is sequentially converted to xanthine and uric acid by xanthine oxidase. A salvage pathway exists whereby hypoxanthine can be converted back to IMP through hypoxanthine phosphoribosyltransferase 1 (HPRT1). Extracellular uric acid levels can also be influenced by modulating its transport rates through importers: uric acid transporter 1 (URAT1) and organic anion transporter 4 (OAT4) and exporters: glucose transporter 9 (glut9), ATP-binding cassette sub-family G member 2 (ABCG2), organic anion transporters 1 and 3 (OAT1 and OAT3), multidrug resistance protein 4 (MRP4) and sodium-phosphate transporter 1 (NPT1).

Fig. 2. *ALDH16A1* gene structure and ALDH16A1 protein in human cell lines. (A) Schematic illustration of the human *ALDH16A1* gene and its spliced variant *ALDH16A1_v2* that lacks an exon (shown by the dashed box and arrow). The location and change caused by the *ALDH16A1**2 SNP associated with gout in the recent Icelandic study [34] is depicted by the blue box and arrow below each figure. PCR forward (F) and reverse (R) primers pair were designed to amplify an 470-bp and 314-bp fragment from *ALDH16A1* and *ALDH16A1_v2* mRNA, respectively. Semi-quantitative PCR analysis of *ALDH16A1* variants (B) and Western blotting analysis (C) of human *ALDH16A1* in human cell lines. These included proximal tubule epithelial (HK-2), hepatoma (HepG-2), embryonic kidney (HEK-293), adenocarcinomic alveolar basal epithelial (A549), neuroblastoma (SH-SY5Y), retinal pigment epithelial (D407) and retinal pigment epithelial (ARPE-19) cell lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 3.

Clustering dendrogram of ALDH16 proteins. Putative ALDH16 amino acid sequences were aligned, and then a dendrogram was generated using a maximum likelihood approach with 1000× bootstrap resampling (labeled at nodes). The ALDH16 family consists so far of 4 subfamilies 16A(mammals), 16B (amphibians and lower animals), 16C (bacteria) and 16D (fish).

Fig. 4. Human ALDH16A1 protein structure and alignment. (A) Predicted conserved domains in the human ALDH16A1 protein by alignment and homology with ALDH2 [33]. In addition the P527R mutation is noted. The two ALDH protein domains (aldehyde DH) are shown as orange bars with a line between them representing the flexible linker. Numbers indicate the amino acid positions of the domains. Additional predicted motifs are shown reflecting four predicted trans-membrane domains (TM), the predicted coiled-coil motif (C-C), as well as the location of the deleted exon in human ALDH16A1v2 (v2 del). (B) Amino acid sequence alignment of several ALDH proteins around Cys-302 of the active site. Note that this region encodes part of the exon missing from ALDH16A1v2. Dashes in sequences correspond to gaps. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 5.

Molecular modeling of human and frog ALDH16 proteins. The secondary structure of human ALDH16A1 (hALDH16A1, left) and frog ALDH16B1 (fALDH16B1, right) are presented as ribbon diagrams (outside figures) with alpha-helices in red, beta-sheets in blue, and coils in green. Space filling models (inner panels) show the substrate-binding pocket (top panels), and cofactor (NAD⁺) binding pocket (bottom panels). Labeled amino acids correspond to residues that are important in substrate or NAD⁺ binding. The substrate binding and NAD⁺ binding pockets are located on opposite faces of the protein, and tunnels between the two is indicated by a white arrow (frog ALDH16B1 only). Residues highlighted in purple (all figures) indicate critical substrate binding amino acids. The location of the deletion that is homologous to Cys-302 in hALDH16A1 is labeled Cys(*). Amino acid residues critical to NAD⁺ binding are highlighted in red (inner panels only, not shown in ribbon diagrams). Details on the amino acid positions compared to hALDH3A1 numbering are provided in Table 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. Homology modeling of two splice variants of human ALDH16A1. Homology models of the long (A) and short (B) forms of human ALDH16A1 consist of globular N-terminal (colored blue in the long form and magenta in the short form) and C-terminal (orange) domains separated by a flexible linker (green). The region of the long form missing from the short form (i.e., Glu253–Pro304) is highlighted in yellow. Pro527 and Pro476 are depicted as yellow spheres in (A) and (B), respectively. Lower figures provide a perspective after 180° rotation of the molecule. Representative energy minimized structures from the predicted protein-protein interaction analysis indicating the capacity for the N- and C-terminal domains of the long (C) and short (D) forms of ALDH16A1 to associate together into a closed conformation. The numbers under each figure are the calculated interaction energies of the respective complexes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 7. Predicted ALDH16A1 and HPRT1 protein complexes. Energy minimized protein complexes of wild-type (A) ALDH16A1 long (N-terminus = blue, linker = green, C-terminus = orange) and (B) ALDH16A1 short (N-terminus = magenta, linker = green, C-terminus = orange) homology models with the crystal structure of human HPRT1 (PDB ID: 1BZY; purple). Molecules of ImmucillinGP, the bound inhibitor co-crystallized with HPRT1, are represented as yellow sticks to indicate the positions of the four active sites of HPRT1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1

Predicted interaction energies between the HPRT1 and the wild-type and mutant ALDH16A1 domains.

	ALDH16A1 long		ALDH16A1 short	
	N-term	C-term	N-term	C-term (P476R)
Interaction energy (kcal/mol)	-19.6	-18.4	-17.8	-17.4
				-16.1

Interaction energies were calculated for HPRT1 and the wild-type (C-term) and mutant (C-term (P527R) or C-term (P476R)) domains of the ALDH16A1 spliced variants (long and short). Mutant and wild-type have the same N-terminus (N-term) domain.

Table 2

Minimized energies of wild-type and mutant ALDH16A1–HPRT1 complexes.

Complex	Energy (kcal/mol)
Long ALDH16A1(WT)–HPRT1	−85,179.2
Long ALDH16A1(P527R)–HPRT1	−85,135.1
Short ALDH16A1(WT)–HPRT1	−82,476.3
Short ALDH16A1(P476R)–HPRT1	−82,414.7

Table 3

Amino acid residues critical to substrate and cofactor binding.

hALDH3A1	fALDH16B1	hALDH16A1
Active site		
Cys-302	Cys-291	Deletion
Glu-268	Glu-257	Ala-272
Asn-169	Asn-160	Thr-176
NAD ⁺ binding site		
Ser-246	Ser-236	Ala-251
Trp-168	Trp-159	Pro-175
Lys-192	Lys-183	Leu-199

Some representative amino acid residues in human ALDH16A1 (hALDH16A1) and frog ALDH16B1 (fALDH16B1) critical to their substrate and cofactor (NAD⁺) binding are provided and related to homologous residues in human ALDH3A1 (hALDH3A1).